DOCUMENT RESUME

ED 026 099 LI 001 282

By-Werner, D. J.; And Others

A Case of Search vs. Experiment. Program of Research on the Management of Research and Development.

Northwestern Univ., Evanston, Ill. Dept. of Industrial Engineering and Management Sciences.
Pub Date Dec 65
Note-7p.
EDRS Price MF-$0.25 HC-$0.45
Descriptors-Experiments, *Information Retrieval, *Information Seeking, *Information Sources, *Librarians, *Researchers, Search Strategies

During a series of operational field experiments to study the information seeking behavior of a sample of medical researchers, a subject reported that, based on his knowledge of information sources at that point it would be less costly to perform an experiment than to attempt to retrieve the piece of information he needed. A professional librarian, working independently, was able to retrieve the same information through use of personal contacts in a pharmaceutical house in a shorter elapsed time than it took to perform the experiment. This case seemed to suggest that some information sources were not well known to the searcher rather than that the information was unavailable, and possibly a greater emphasis needs to be placed upon informing the users about current information resources. (JB)

ED 001282

ED 026099

ERIC/CLIS
FEB 25 '69
UNIVERSITY OF MINNESOTA

PROGRAM OF RESEARCH ON THE MANAGEMENT

OF RESEARCH AND DEVELOPMENT

Department of Industrial Engineering and Management Sciences
The Technological Institute
Northwestern University
Evanston, Illinois

A CASE OF SEARCH VS. EXPERIMENT

D. J. Werner, G. J. Rath and A. H. Rubenstein

December, 1965

ABSTRACT

During a real-time study of the information-seeking behavior of medical researchers, a subject reported that it would be easier to perform an experiment than to attempt to retrieve the piece of information he needed. A professional librarian, working independently, was able to retrieve the same information in a shorter elapsed time than it took to perform the experiment.

ED 001282

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

65/33

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

A Case of Search Vs. Experiment

The growing number of people concerned with new information retrieval methods reflects the fact that the scientist faces great difficulties in finding the information he needs. Indeed, the comment is frequently made that in many cases the researcher finds duplication of an experiment easier, cheaper, and more efficient than attempts to retrieve the needed information. A strong argument for improved information retrieval systems is, therefore, the possibility of reducing this duplication by making retrieval the more desirable alternative. In effect, this is expected to help increase the total output of the scientific community. Unquestionably, duplication has a real function in some areas. On the other hand, constant repetition of routine work serves no real purpose. if it can be avoided.

One current project of the Program of Research on the Management of Research and Development[1] at Northwestern University involves studying the information-seeking behavior of a sample of medical researchers through a series of operational field experiments. During one such study[2], an event related to search-vs.-experiment occurred. One of the researchers working in a basic research department of a major medical research organization, reported (through our remote data collection system)[3] that he was going to repeat an experiment because he did not think it would be worth the "time and trouble" to try to find the information he needed.

His question was: Can substance X be successfully treated by process Y? He did know that other less refined substances had been handled by this process, but he did not know if such a highly refined material, as X, could also be treated successfully. Since he did not know where the information was readily available, he repeated the experiment.

This experiment required four hours of a technician's time spread over two days because of a required waiting period. Approximately twenty minutes of the researcher's time was needed to arrange and explain the experiment to the technician.

In his decision to repeat the experiment, Dr. B. indicated that with the present state of knowledge, repetition would be less costly than retrieval. Since the information was not urgently needed, he was merely minimizing his "cost". The technicians available are not trained to do searching, but they can perform such experiments. Dr. B. himself, however, would have had to do the searching.

The situation can be visualized with the aid of the following model:

S_4 = The total universe of information sources potentially of use to the researcher including both formal and informal sources.

S_3 = All the sources of information of which the researcher is aware.

S_2 = All of those sources of information which the researcher is capable of using at a given moment due to restrictions of finance, distance, time, etc.

S_1 = All of those sources which the researcher currently is using either on a regular or an irregular basis.

Dr. B. was in effect indicating that as far as he knew, the source of the information did not exist in his S_3. For him to retrieve the information, he would have had to use a source presently in S_4-S_3 at some undetermined cost, where the cost would include time, money, effort, etc.

Two days after the event occurred, Mrs. A. - an information specialist - was assigned full-time to Dr. B's group, to study the effect of her services on the information-seeking behavior of the researchers. This provided an excellent opportunity to determine the difficulty which someone acquainted with a broader range of information sources would have in answering the question which Dr. B. had encountered. Mrs. A. was given all the information which we obtained from Dr. B. when he described the problem, including information about the other uses of the process. She was instructed to keep an accurate account of the time spent and the sources used in finding this information. Her results were as follows:

Source Used	Time	Results
John Crerar Library* Card Catalog	10 min.	Negative
Open shelves at Crerar Library	15 min.	Negative
Called a friend at a chemical company who was familiar with processes related to Y.	5 min.	Referred to a Dr. C. at another hospital
Called Dr. C., but spoke to Dr. D.	3 min.	Referred to another doctor at the same hospital who recommended Mr. P. at a pharmaceutical house
Called Mr. P. at the pharmaceutical house.	3 min.	Yes, X can be treated by process Y. He gave the name of two pharmaceutical houses which had the requested information.

*A technical-medical library located in Chicago.

The time taken by Mrs. A. was thirty-six minutes. Of this, 70% was spent checking written sources which were of no assistance. She indicated she would not have used them had she not been at Crerar Library when she began the search, because at the hospital she would have had immediate access to a telephone and would have tried the verbal sources first. Thus it took about twice as long for her to find the material as it did for Dr. B. to instruct his technician to perform the experiment. However, had she been available at the hospital, she would have spent only about half of the time spent by Dr. B.

Given these facts, the questions become: Should he have repeated the experiment? What difference in cost would there have been had he done the search, or if Mrs. A. had been available to help in the search? The first of these is difficult to answer without an extensive "cost-effectiveness" study. At the time he needed the information, he perceived a repetition to be the best path to the information. In addition, the brevity of Mrs. A's search hinged upon having a personal contact who was the beginning of the series of connections leading to the information. Dr. B., however, perceived the expected value of the search to be small enough not to justify the search. Had he searched unsucessfully, he would have still had to repeat the experiment. Since he had the technician's service, he chose what seemed the best method. On the other hand, had Mrs. A. been available at the time he might have found the use of her services to be a better alternative.

As indicated, the searcher is presumably maximizing his expected values. At least in some way he evaluates the cost of alternatives

and selects what appears to be best, based on his knowledge at that point. Unfortunately, the knowledge he has may not always be what is necessary to make the "correct decision". In this case, twenty minutes of his time and four hours of technician time seemed a desirable alternative to an undefined searching period with an unknown and probably low probability of success. Had the use of personal contacts in a pharmaceutical house been among those available in his S_1 or S_2, then the search may have been quite different, perhaps very similar to that of Mrs. A.

While no generalization can be made from one instance, this case seems to suggest that some information sources are not well known to the searcher rather than that the information is unavailable. (This lack of acquaintance with sources occurred in other instances in the experiment.) This seems to be a significant but quite often overlooked point. A great emphasis is currently being placed upon developing complex new systems to deal with the information problem. While some of these new systems are undoubtedly useful and necessary, quite possibly a greater emphasis needs to be placed upon informing the users about current information resources[4]. Quite possibly, many of the new systems that are developed will proceed along the same path as many in the past. They are placed in the S_4 but never reach the S_1 of the individual scientist. Obviously, more than a sales campaign is needed to acquaint the researcher with information sources and to <u>lead him to use them</u>. If the researcher could be better informed of and instructed in the general use of information systems at some point in his career, the existing systems might function more smoothly and perform more effectively for the individual researcher.

REFERENCES AND NOTES

1. Albert H. Rubenstein, "A Program of Research on the Research and Development Process," IEEE Transactions on Engineering Management, September, 1964, No. 3, Volume EM-11, pp. 103-112.

2. David J. Werner, "A Study of the Information-Seeking Behavior of Medical Researchers," Master's Thesis, Department of Industrial Engineering and Management Sciences, Northwestern University, December, 1965.

3. Gustave J. Rath, "Initial Steps Toward Studies in the Information Behavior of Scientists and Engineers," Second Congress of the Information Systems Sciences, Edited by J. Speigel, Spartan Press, 1965.

4. Albert H. Rubenstein, "Timing and Form of Researcher's Needs for Information," Journal of Chemical Documentation, 2, 28 (1962).

5. This research is supported by Grant NSG-495, from the National Aeronautics and Space Administration.